

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許出願公告番号

特公平6-87241

(24)(44)公告日 平成6年(1994)11月2日

(51)Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 15/20	5 7 0 Z	7315-5L		
12/00	5 2 0 E	8944-5B		
15/40	5 0 0 Z	9194-5L		

請求項の数 3 (全 9 頁)

(21)出願番号	特願平2-276572	(71)出願人	999999999 インターナショナル・ビジネス・マシーンズ・コーポレーション アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)
(22)出願日	平成2年(1990)10月17日	(72)発明者	マービン・エル・ウィリアムズ アメリカ合衆国テキサス州レウイスビル、 セトラーズ・ウェイ1152番地
(65)公開番号	特開平3-191467	(74)代理人	弁理士 頓宮 孝一 (外1名)
(43)公開日	平成3年(1991)8月21日		
(31)優先権主張番号	4 5 4 7 9 7		
(32)優先日	1989年12月19日		
(33)優先権主張国	米国 (U S)		
		審査官	高橋 英生
		(56)参考文献	特開 昭58-158769 (JP, A)

(54)【発明の名称】 文書属性の識別方法

【特許請求の範囲】

【請求項1】複数の文書が記憶され、且つそれらの文書を結合してある文書関係を形成することが可能な共用ライブラリにおける文書属性の識別方法であって、上記複数の文書のうち上記文書関係に対応する少なくとも2つの文書を識別するステップと、上記の少なくとも2つの文書の文書属性を含むベクトル関係特性オブジェクトを生成するステップと、よりなることを特徴とする識別方法。

【請求項2】上記のベクトル関係特性オブジェクトを生成するステップが、上記の少なくとも2つの文書の中の1つが現在のベクトル関係特性オブジェクト中に既に存在するか否かを判断するステップと、上記の少なくとも2つの文書の上記文書属性を上記の現

在のベクトル関係特性オブジェクト中に記憶するステップと、を含むことを特徴とする請求項(1)記載の識別方法。

【請求項3】複数の文書が共用ライブラリに記憶され、且つ上記複数の文書にある文書関係を形成させることが可能である文書管理システムにおける文書属性の識別方法であって、上記複数の文書のうち少なくとも2つの文書間に上記文書関係を生成するステップと、上記の少なくとも2つの文書の文書属性を識別するステップと、上記の少なくとも2つの文書の上記文書属性を関係特性オブジェクト中に記憶するステップと、よりなることを特徴とする識別方法。

【発明の詳細な説明】

A. 産業上の利用分野

この発明は情報処理システムにおける文書管理方法に関し、特に共用ライブラリにおいて同じ属性を有する文書を迅速に識別する方法に関する。

B. 従来の技術

電子オフィスにおいては、文書を電子的に記憶し、探索し、検索するのに、これらの処理を行い易くすることを目的とした幾つかの関係に従い文書をグループ化することが行われている。この必要を満たすための技術の代表的なものとして、文書ライブラリを用いるものがある。現行の文書ライブラリによれば、手作業による書類のファイリング及び検索に類似した機能が行われるが、その煩雑さは手作業に比べてはるかに軽減されている。文書を電子的に記憶し、探索し、検索するという必要を満たす先行技術ストラクチャの一つに、インターナショナル・ビジネス・マシーンズ・コーポレーションにより販売されているオフィス・システムの一部をなす文書交換アーキテクチャ (DIA) がある。DIAの文書ストラクチャでは、ライブラリにファイルされる各文書毎に一組の記述子セットが用いられる。これらの記述子は文書プロフィールに入れられ、文書と共にファイルされる。文書プロフィールは、ファイルする際の文書名、著者、対象課題、文書ライブラリにファイルした日付などのような文書内容を識別するためのパラメータを内容とする。文書プロフィールは文書ライブラリ中の文書を探索するのに用いられる。例えば、ユーザーは2日間にライブラリが受け入れたある著者による特定主題に関する全ての文書について照会することができる。その探索が完了すると同時に、ユーザーは探索基準に合致する文書のリストを与えられることになる。ユーザーないしはエンドユーザーは、データ処理や情報交換のためにライブラリを利用する人、装置、プログラム、あるいはコンピュータ・システム等である。

特定の探索基準に基づいて文書を探索する他、DIA文書ライブラリ・サービスには、文書間に適宜の関係を確立するための手段が講じられている。

このような文書間の関係 (以下文書関係とする) の一例として、フォルダー文書概念がある。フォルダー文書はユーザーが文書を線形構造または階層構造にグループ化する場合に生成される。文書は、各文書セットまたは集合を特定フォルダー内のユーザー指定場所に入れることによりそれらの文書を編成するとき線形構造をとる。また、文書は、フォルダーである文書をネスティングする、より平易に言うと、それ自体フォルダーである文書をやはりフォルダーである他の文書の中に入れることによって階層構造をとることができる。フォルダー中の各文書は、他の文書との関係に関わらず、記述子、アクセス制御、及び分素内容を含めてそれぞれ別個の文書定義を有する。従って、ある特定の関係に含まれる1つの文書に対するアクセスは、それと同じ関係に含まれる他の

文書へのアクセスに関しては何の意味も持たない。DIA文書ライブラリ中の全ての文書は、各文書に付随して文書関係を記述する文書関係オブジェクトと呼ばれるライブラリ・モデル・オブジェクトを有する。そのため、あるフォルダーをアクセスすることが必要なときは、フォルダー中の各文書の各文書関係オブジェクトをアクセスして文書間の全ての関係を判別しなければならない。

DIAライブラリにおいて可能な文書関係のもう一つの例として、ステープル関係と呼ばれるものがある。このステープル関係によれば、ユーザーは2つの文書を互にくっ付けることができる。ステープル関係は、文書間に緊密に結合された一対一の指示された文書関係を与える。さらに、ステープル関係により結合された文書はフォルダー文書中に入れることができる。各文書の文書関係オブジェクトは、フォルダー文書に関連して前に述べたように、フォルダー文書間の関係を記述すると共に、ステープル関係により結合された文書間の関係を記述しなければならない。このように、文書関係中の各文書には、言語、文書形式 (document type)、文字セット、及び文書自体の特性を記述するその他のプロフィール情報のような各文書独特の属性が入っている。例えば、ある関係中の全ての文書がフランス語で書かれているとすると、それらの各文書には文書がフランス語で書かれているということを指示するプロフィール情報が入っている。

C. 発明が解決しようとする課題

上記のような従来技術にあっては、ユーザーが文書関係によって定義される文書をアクセスしようとする場合、その関係に含まれる全ての文書の属性を判別するのに、各文書のプロフィールを読み、その内容を解釈しなければならない。その結果、一旦それらの属性を指定してから、それら特性の属性を含む関係内の全ての文書を識別する手段が必要となる。

D. 課題を解決するための手段

本発明は、複数の文書中に重複して見られる属性を指定する際の文書管理システムの必要記憶量及び複雑性を減じる方法に関するものである。文書管理システムでは、言語、文書形式、及び文字セットのような属性を文書内容と共に記憶する。個々の文書を他の文書と結合してフォルダーやステープル関係文書のような関係を生成するためには、各文書に照会してそれぞれの属性を判別しなければならない。そこで、本発明は、2つ以上の文書間にある関係が定義される時生成されるベクトル関係特性オブジェクトを開示するものである。このベクトル関係特性オブジェクトは、その関係中の少なくとも1つの文書に関する特定の属性を識別するためのフィールドを有する。このベクトル関係特性オブジェクトは、文書アクセス機において、その関係中の個々の各文書をそれぞれの属性に関して走査する必要なくそれらの文書属性を迅速に判別するのに利用することができる。

E. 実施例

第1図は、情報処理システムにおいて文書を管理するための文書管理システムの一例を示している。ユーザー20はライブラリ28に文書を記憶し、且つライブラリ28の文書を検索することができる。ライブラリ28は同時に他のユーザーもアクセスすることができ、従って共用ライブラリをなしている。ライブラリ28と異なり、ユーザーのパーソナル文書あるいは専用文書はローカル文書記憶装置24に記憶される。このローカル文書記憶装置24は、通常、他のユーザーと共用されない。ユーザー20はマネージャ（管理プログラム）／リクエスト22を介して自己のローカル文書記憶装置24にアクセスする。マネージャ／リクエスト22は、ユーザーのライブラリ28に対するアクセスを制御するライブラリ・サーバ26とのインターフェースとしての機能を有する。ユーザー20がライブラリ28に文書をファイルする時、ライブラリ・サーバ26はライブラリ28に記憶中の情報の内容を記述するパラメータまたは記述子を構成しなければならない。

第2図は、第1図のライブラリ28に記憶される情報の基本文書モデルを示す。この文書モデルはライブラリ・サーバ26によって生成され、各文書と共に記憶される。DIAライブラリ・サービス・アーキテクチャをサポートするには、この基本文書モデルを明確な形で実施する必要はないが、設計をこれらのモデルまたはその部分集合に合わせてマッピングすることは必要である。

ユーザー・プロフィール・オブジェクト（UPO）36はDIA文書モデルの一部ではなく、DIA文書モデルによって引用されるオブジェクトである。ユーザー・プロフィール・オブジェクト36は、これによって表されるユーザーがオフィス・システム・ネットワークのログイン（login）された加入者である時生成される。ユーザー・プロフィール・オブジェクト36はユーザーの識別情報を与えると共に、ユーザーに対して許可された別名、サービスのようなユーザーに関する情報、デフォルト（省略値）解釈情報（default accounting information）、及びその他のユーザー固有の情報を含む。文書モデル・オブジェクト30はDIAモデルの心臓部をなし、論理的にはある文書が初めて文書ライブラリにファイルされる時生成される最初のオブジェクトである。このオブジェクトには各特定文書の所有権及び属性に関する情報が入っている。より詳しく言うと、文書モデル・オブジェクト30は、文書が編集可能か編集可能でないかというような文書インスタンス属性、最大版数、編集することのできない文書をユーザーが編集しようとした時とるべき処置等を内容とする。さらに、文書モデル・オブジェクト30には、下記の任意の情報を入れることができる。

1) 更新のためにライブラリから文書を出してチェックした者をユーザーに指示する文書レベル・ロッキング情報。

2) 文書内容がライブラリ・サーバの直接制御から除

外されたと言う指示。

3) ライブラリ・アドミニストレータ（library administrator）に対する文書の一部をライブラリ・サーバより除くかまたはライブラリ・サーバへ戻すようにと言う指示。

4) 時刻・日付情報がライブラリ・サーバの直接制御から除外されたと言う指示。

5) ライブラリ・サーバの直接制御から除外された情報の記憶場所。

6) アーカイバル（長期保存）目的の任意情報。

アクセス制御モデル・オブジェクト（ACMO）32は、ある文書が初めてDIAライブラリにファイルされる時生成される。アクセス制御モデル・オブジェクト32の主たる目的は、文書に対する非オーナー・アクセスを判別するのに利用することができるよう、情報を整理・統合（consolidate）することにある。このオブジェクトは、その情報が誰でもアクセスすることができるもの（公衆用）であるか、少数の明確に指定されたユーザーにのみアクセスが許可されるもの（専用）であるか、あるいはその情報が他のユーザーと共用されるものであるかというようなアクセス制御情報を含む。アクセス制御モデル・オブジェクト32には、文書の保存及び廃棄を管理する情報も含まれる。

文書履歴記録オブジェクト（DHL0）34は、文書がライブラリにファイルされ、ユーザーがその文書に対する様々な活動を記録したいと欲する場合に任意に生成される。例えば、ユーザーが文書が読まれた回数及び文書を読んだ者を記録したいと考える場合などである。

文書関係オブジェクト（DRO）42はある文書が初めてライブラリにファイルされる時生成される。その目的は、ある文書とこれと関係付けられた他の文書またはこれと同じグループの他の文書との論理関係を記述することにある。例えば、DIAアーキテクチャにおいては、他の文書を含むフォルダー文書を生成することができる。このような関係が存在する場合、フォルダーに入っている各文書は文書関係オブジェクト42中にライブラリ割当て文書名（LADN）と呼ばれるポインタ・エントリを有する。版制御オブジェクト（VCO）40は、ある文書が初めてライブラリにファイルされ、且つ文書の単一の名前付き版よりなるいくつかのオブジェクトに関する情報を含む場合に生成される。このオブジェクトは、版ネーミング、版レベルのロック、及びその他の版関係レベルのプロセス制御のためのスペースを提供する。

プロフィール内容オブジェクト（PCO）44は、ある文書が初めてライブラリにファイルされ、且つユーザーが性能またはその他の理由のためにサブオブジェクトを生成したいと欲する場合に生成される。このプロフィール内容オブジェクト44はサブオブジェクトに関するプロフィール情報用の保管場所（repository）である。

文書内容オブジェクト（DCO）46は、ある文書が初めて

ライブラリにファイルされる時生成され、文書内容の記憶のために用いられる。さらに、文書内容オブジェクト46は種々の測定単位における実際の文書サイズに関する情報をセーブするための記憶場所としても用いられる。探索インデックス・オブジェクト(SIO)48は、文書内の探索に用いられるエントリを内容とする。これらのエントリは他のオブジェクトに対する下記の一連の動作の結果としてSIOに入れられる。まず、基本をなす文書モデル・オブジェクト30がFILEコマンドの処理の一部として生成される。次に、ライブラリ・サーバがプロフィール内容オブジェクト44、文書関係オブジェクト42、及びアクセス制御モデル・オブジェクト32を走査して、パラメトリックのSEARCHをサポートするために用いる項(terms)を探す。各探索項が識別される毎に、名称にパラメトリック探索項の値及びセマンティクスが含まれる探索インデックス・オブジェクト48にエントリがなされる。ライブラリ・サーバが上記オブジェクトを走査する時SIO48がない場合には、SIO48が1つ生成され、あたかも探索インデックス・オブジェクト48が常時存在するかの如く、そのSIOに対してエントリがなされる。

逆探索インデックス・オブジェクト(RSIO)38は、文書をDELETEコマンドによってライブラリから取り除く時探索インデックス・オブジェクト48の除去をサポートするために存在する。パラメトリック探索項に関するRSIO38へのエントリは、探索インデックス・オブジェクト48に対するエントリと同時になされる。

第3図はフォルダーの代表的な文書関係を示す。文書A(56)及び文書B(54)はフォルダー文書1(50)に含まれる個々の項目を表す。文書A(56)は文書C(58)と共にフォルダー文書2(52)にも入っている。フォルダー文書1(50)及びフォルダー文書2(52)は例えば個々のパーソナル・ジャケットであり、文書A(56)は例えば個人のポリシーを説明した文書である。文書A(56)の実コピーは1枚だけ必要であり、そのコピーは全てのパーソナル・ジャケットにわたって常に一貫性を保持しなければならない。文書とフォルダーとの関係は文書関係オブジェクト(DRO)に入っている。各フォルダーへのアクセスのためには、ライブラリ・サーバが各文書の文書関係オブジェクトをアクセスしてその関係を判別する必要がある。

第4図には、フォルダー概念における文書の拡張においてはフォルダー中にフォルダーが含まれると言うことが図示されている。文書A(56)及び文書B(54)はフォルダー文書2(52)中の項目である。次に、フォルダー文書2(52)はフォルダー文書1(50)に入れられる。例えば、フォルダー文書1(50)は社内世論調査報告であり、文書2(52)はオースチン(Austin)の調査報告である。文書A(56)及びB(54)はオースチンの異なる2つの地域からの調査報告であってもよい。この場合も、ライブラリ・サーバがこのような関係によって示さ

れる文書をアクセスする毎に、各項目の文書関係オブジェクトを吟味しなければならない。

第5図には、1つの文書が他の文書に結合されるもう一つの文書関係が示されている。文書1(60)は文書3(64)に結合された文書2(62)に結合されている。このステープル関係によれば、ユーザーは2つの文書をくっ付けることができる。各文書の文書関係オブジェクト(DRO)には文書関係の現在の状態が入っている。

第6図には、ステープル関係とフォルダー関係の他の組合せの例が示されている。文書1(60)と文書2(62)は互いに結合して、文書4(66)と共にフォルダー文書1(50)に入れることができる。同様に、文書3(64)も文書2(62)と結合して、文書フォルダー2(52)に入れることができる。文書5(68)もフォルダー文書2(52)中に入っている。フォルダー中の各文書には、言語、文書形式、文字の種類、文字セット及び文書自体の特性を記述するその他のプロフィール情報等、各文書独特の属性が含まれている。フォルダー中の文書をアクセスするためには、ライブラリ・サーバは各文書のプロフィールを読み、解釈してその文書関係中の全ての文書の属性を判別しなければならない。例えば、その文書関係中の全ての文書がフランス語で書かれているとすると、各文書はその文書がフランス語で書かれていると言うことを指示するプロフィール情報を有するはずである。本発明は、各文書毎にこのような属性を入れ、記憶しなければならないと言う問題を、「ベクトル関係特性オブジェクト(VRCO)」を生成することによって解消するものである。

第7図は、本発明のベクトル関係特性オブジェクト(VRCO)70を用いたDIA文書ライブラリ・サーバ用の基本文書モデルを示す。VRCO70は、ある関係内の少なくとも1つの文書について存在する特性属性を識別するためのフィールドを有する。これらのフィールドの後には、それらの特性属性を有する文書を一意に識別する識別子が続く。

第8図及び第9図はVRCO70におけるフィールド構成の典型例を示す。第8図において、文書1、2及び3がいずれも属性Xを持ち、文書1、2及び6が属性Yを持つとすると、ベクトル関係特性オブジェクト(VRCO)は図示のようなフィールドを有する。属性X及びYに関しては、その情報が既にVRCOに入れられているから、各文書のプロフィールで指定する必要はない。新しい属性の出現はそれまでの属性を有する全ての文書の終わりを示す。

本発明によれば、ライブラリ・サーバは、アプリケーションよりVRCOを読むことによって要求された時、相互に関係付けられている文書の属性を迅速にアクセスすることができる。このように、文書相互の関係は、それらの関係がVRCOに記憶されている場合迅速にアクセスすることができる。そこで、迅速なアクセスを必要とする属性

／特性をVRCOに入れると言う構想を、以下の例によってさらに明確に説明する。まず、アプリケーションはある文書の内容に関して特定の文字セットしか理解することができないものと仮定する。すると、アプリケーションは、その特定の文字セットを持つ文書を探索しつつVRCOを迅速に走査することができるはずである。次に、アプリケーションはそれらの文書のみを表示のためにユーザーに提示し、アプリケーションがサポートしていない文書は無視または削除することができよう。従って、本発明はアプリケーションをしてアプリケーションがサポートし得ない特性を持つ文書用のリソース（資源）を迅速に解放することを可能ならしめるものである。

ベクトル関係特性オブジェクト（VRCO）には、ベクトルが属性以外または文書識別子以外にベースする（based off of）べきものであるかどうかを指定するために、VECTOR-BASE値を入れることもできる。第9図には、VRCOにおけるこれらのフィールドの構成がより明確に示されている。VECTOR-BASEフィールドには、そのベクトルが文書識別子または文書属性以外にベースされているか否かを示す値が入る。文書1は属性X、Y、及びZを有し、文書2はY属性を有する。このようにベクトルを文書以外にベースすることも可能なようにすることによって、第9図に示すように文書数が定義済み属性の数との関係において比較的多い場合、記憶空間を節約することができる。

第10図は、ベクトル関係特性オブジェクト（VRCO）の利用形態を示すフローチャートである。ステップ71において、ユーザーは文書Aと文書Bとの間にある関係を生成することを決定する。すると、まず文書AがVRCOを有するか否かがステップ72で判断される。文書AがVRCOを持っていれば、次にステップ74で文書BがVRCOを有するか否かが判断される。これらの文書がいずれも既にベクトル関係特性オブジェクトを有する場合は、ステップ76で文書A中にない文書Bの属性があればそれらの属性を識別する。VRCOの分析は、ステップ78でこれらの文書間の新しい関係を反映するよう文書AのVRCOを修正することにより完了する。文書間の関係を定義するVRCOは1つしか存在しない。

ユーザーが文書Aと文書Bの間にある関係に形成しようと決定した後、文書AについてVRCOがないことが判明した場合は、ステップ84で文書BについてはVRCOが存在するか否かを判断する必要がある。その結果、文書BがVRCOを有する場合は、ステップ86で文書Aよりプロフィール情報が抽出され、文書BのVRCOに入れられる。文書Aの属性はステップ88でも識別され、文書BのVRCOに入れられる。最後に、ステップ90で、文書BのVRCOが文書Aの属性を反映するよう修正される。この方法では、各文書関係毎に1つのVRCOが得られる。文書関係の中で文書が何件生成されるかに関わらず、その文書関係に対してはVRCOは1つしか存在しない。これによって、文書の属

性を保存するために必要な記憶量が著しく少なくなる。同様に、文書AについてはVRCOがあり、文書BについてはVRCOが存在しない場合は、ステップ82で文書Bよりプロフィール情報が抽出される。ステップ80で文書Bの属性が識別され、ステップ78で文書AのVRCOがこれらの文書間の新しい関係を反映するよう修正される。

最後に、どちらの文書にも各々に関連するVRCOがない場合は、ステップ73で文書A用のVRCOを生成する必要がある。この場合、ステップ82で文書Bよりプロフィール情報が抽出され、そのVRCOに入れられる。ステップ80で文書Bの属性が識別され、VRCOに入れられる。ステップ78において、文書Aと文書Bの間の新しい関係を反映するよう文書AのVRCOが修正される。

要約すると、本発明は複数の文書が独特の文書関係に結合される場合に文書属性を識別する方法を開示するものである。これらの文書関係には、フォルダー、ステープル文書、フォルダーとステープル概念の種々の組合せ、あるいはある関係内の文書が少なくとも1つの類似の属性を含むような何らかの関係が含まれる。ベクトル関係特性オブジェクト（VRCO）は、文書交換アーキテクチャ（DIA）において実施されているような基本文書モデルを形成するオブジェクトの集合に付加される。ベクトル関係特性オブジェクトは文書アクセス機構で用いることができ、ある関係中の各文書毎に1つだけ属性を指定するフィールドを有する。このVRCOは、文書関係が生成された時1回だけ記憶を行うようにすることにより必要記憶量を減少させることができる。VRCOは文書関係中の少なくとも1つの文書について存在する特定の属性を識別するためのフィールドを有する。このフィールドの後には、その特定の属性を持つ文書を一意に識別する識別子が続く。この属性は単一のオブジェクトに1回だけ記憶されるので、これを迅速にアクセスすることにより、特定の文書についてある属性が存在するか否かを判断することができる。

F. 発明の効果

本発明の方法によれば、複数の文書に重複して現れる属性を指定する際の文書管理システムの必要記憶量及び複雑さを著しく減じることが可能である。

【図面の簡単な説明】

第1図は本発明を実施することが可能な文書管理システムのブロック図、

第2図は第1図に示すシステムのライブラリにファイルされる各文書毎に生成される基本文書モデルのブロック図、

第3図はフォルダー文書の概念を図解した文書関係の一例のブロック図、

第4図はフォルダーを別のフォルダー中に入れたフォルダー関係を示すブロック図、

第5図はステープル関係を示すブロック図、

第6図はフォルダー中でステープル関係に結合された文

書のブロック図、

第7図は第2図の基本文書モデルに本発明の方法によるベクトル関係特性オブジェクトを付加した構成を示すブロック図、

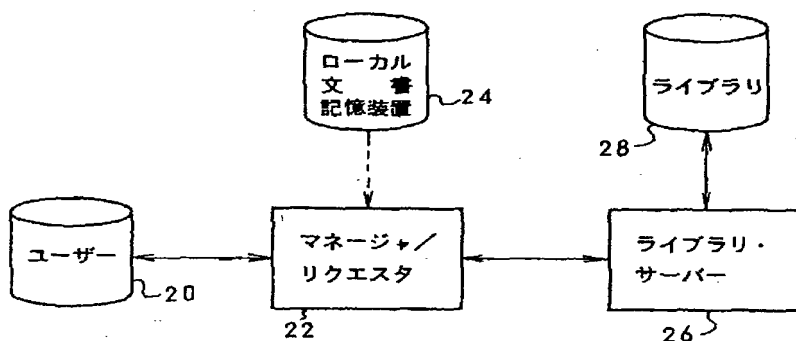
第8図及び第9図はベクトル関係特性オブジェクトにおけるフィールドの構成を示す説明図、

第10図は第1図の文書管理システムにおけるベクトル関係特性オブジェクトの利用形態を示すフローチャートである。

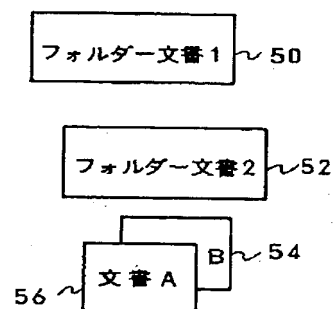
20……ユーザー、22……マネージャ／リクエスタ、26……ライブラリ・サーバ、28……ライブラリ、30……文書モデルオブジェクト、32……アクセス制御モデルオブジェクト、36……ユーザー・プロフィール・オブジェクト、

42……文書関係オブジェクト、44……プロフィール内容オブジェクト、46……文書内容オブジェクト、70……ベクトル関係特性オブジェクト。

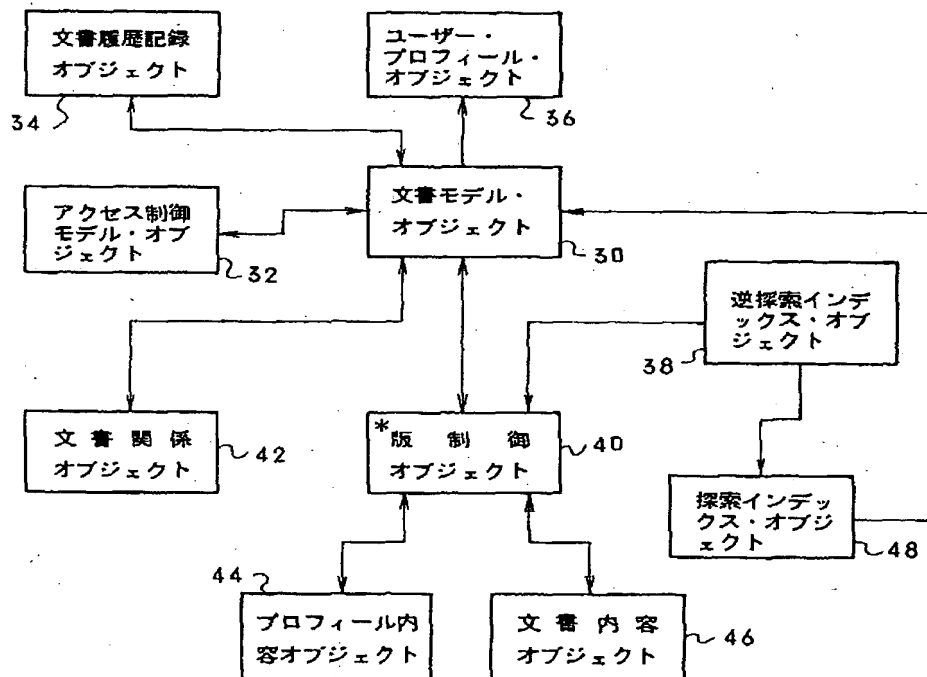
【第1図】



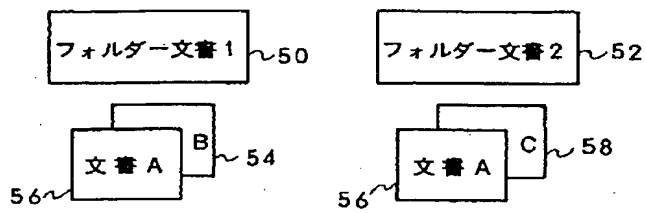
【第4図】



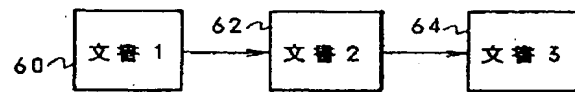
【第2図】



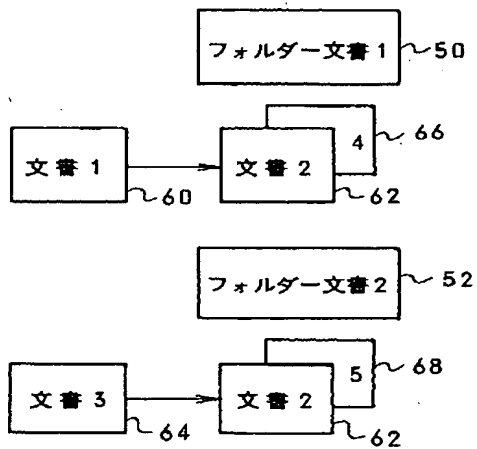
【第3図】



【第5図】



【第6図】



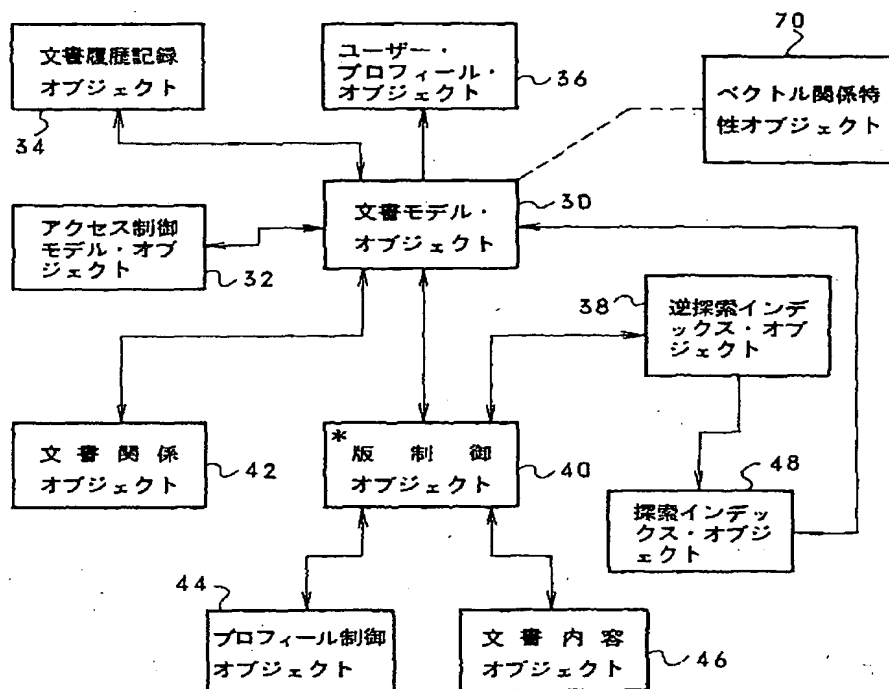
【第8図】

属性 X	文書 1	文書 2	文書 3	属性 Y	文書 1	文書 2	文書 6	...
------	------	------	------	------	------	------	------	-----

【第9図】

文書 1	属性 X	属性 Y	属性 Z	文書 2	属性 Y
------	------	------	------	------	------

【第7図】



【第10図】

